

Amendment to the Claims:

This listing of claims will replace all prior versions, and listing, of claims in the application.

Listing of Claims:

1. (currently amended) A computer implemented method of merging with a base assembly of molecules one or more additional assemblies of molecules, similar molecules in the assemblies having previously been identified and removed using a molecular structural descriptor validated as possessing a neighborhood property, comprising the steps of:

- a) using a molecular structural descriptor, validated as possessing a neighborhood property, which is appropriate to whole molecules, characterizing all the molecules in the base assembly of molecules and in the assembly of molecules to be merged;
- b) calculating the molecular structural distance between every molecule in the base assembly to every molecule in the assembly to be merged;
- c) while there are still molecules in the assembly to be merged which have not been tested, selecting a molecule from the assembly to be merged;
- d) determining whether the molecular structural distance between the selected molecule and every molecule in the base assembly is within the neighborhood distance of the molecular structural descriptor;
- e) selecting for inclusion in the merged assemblies only those molecules identified in step d as having molecular structural distances greater than the neighborhood

distance.

- f) repeating step c through step e until all molecules in the assembly to be merged have been tested; and
- g) repeating step a through step f for each additional assembly to be merged; and
- h) outputting the merged assembly of molecules -

wherein the merged assembly of molecules has maximum diversity and minimum redundancy between and among its members.

2. (previously presented) The method of claim 1 in which the molecular structural descriptor, validated as possessing a neighborhood property, appropriate to whole molecules is the Tanimoto similarity coefficient.

3. (currently amended) A computer implemented method of merging with a base assembly of molecules one or more additional assemblies of molecules, similar molecules in one or more of the assemblies having not previously been identified and removed using a molecular structural descriptor, validated as possessing a neighborhood property, comprising the steps of:

- a) selecting subsets of each assembly by:
 - (1) selecting a molecule within each assembly;
 - (2) using a molecular structural descriptor, validated as possessing a neighborhood property, appropriate to whole molecules, calculating the descriptor distance between the selected molecule and all molecules within the assembly;
 - (3) determining the shortest distance between the selected molecule and all

- molecules previously selected for the subset;
- (4) selecting for inclusion in the subset the molecule whose shortest descriptor distance from the previously selected molecules is the largest and is greater than the neighborhood distance of the descriptor;
 - (5) repeating steps (1) through (4) until the largest shortest difference between molecules is less than the neighborhood distance of the descriptor; and
 - (6) repeating steps (1) through (5) for each assembly;
- b) using a molecular structural descriptor, validated as possessing a neighborhood property, which is appropriate to whole molecules, characterizing all the molecules in the base assembly of molecules and in the assembly of molecules to be merged;
 - c) calculating the molecular structural distance between every molecule in the base assembly to every molecule in the assembly to be merged;
 - d) while there are still molecules in the assembly to be merged which have not been tested, selecting a molecule from the assembly to be merged;
 - e) determining whether the molecular structural distance between the selected molecule and every molecule in the base assembly is within the neighborhood distance of the molecular structural descriptor;
 - f) selecting for inclusion in the merged assemblies only those molecules identified in step e as having molecular structural distances greater than the neighborhood distance.

- g) repeating step d through step f until all molecules in the assembly to be merged have been tested; and
- h) repeating step b through step g for each additional assembly to be merged; and
- i) outputting the merged assembly of molecules -

wherein the merged assembly of molecules has maximum diversity and minimum redundancy between and among its members.